

Abstract of CN 1400541 A

The invention discloses a multi-media real time control lessons giving system based on IP network, it is made up of lessons giving end, listening end, class service center and multipoint control unit(MCU); it realize a live telecast of teaching through multi-media transcribing the teaching spot and network transmission, through multi-media and multi-mode intercommunication network the teachers and the students natural blackboard-writing teaching by the teacher, teaching content retrieving, share of applied procedure, teaching software skimming, electronic white board, real time control teaching software and the class management function, it broads the teaching scale.



[12] 发明专利申请公开说明书

[21] 申请号 02139372.9

[43] 公开日: 2003 年 3 月 5 日

[11] 公开号 CN 1400541A

[22] 申请日 2002.8.20 [21] 申请号 02139372.9

[71] 申请人 西安交通大学

地址 712100 陕西省西安市咸宁西路 28 号

[72] 发明人 郑庆华 刘 均 李 洋

[74] 专利代理机构 西安通大专利代理有限责任公
司

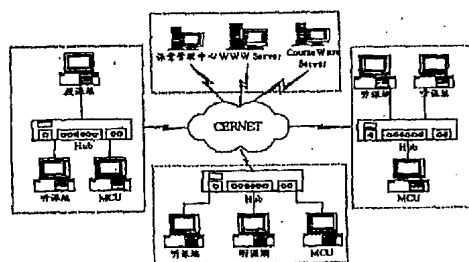
代理人 李郑建

权利要求书 3 页 说明书 14 页 附图 6 页

[54] 发明名称 基于 IP 网的多媒体实时授课系统

[57] 摘要

本发明公开了一种基于 IP 网的多媒体实时授课系统, 其由授课端、听课端、课堂服务中心以及多点控制单元 (Multipoint Control Unit, MCU) 四个部分组成; 它通过教学现场的多媒体录制和网络传输, 实现教学现场的直播, 并通过师生多媒体多模式交互、教师自然板书授课、教学内容检索、应用程序共享、课件同步浏览、电子白板、课件实时录制、课堂管理等功能, 解决传统课堂教学在时间和空间上的制约问题, 大大扩展了教学规模, 实现名师授课及教育资源的共享。



ISSN 1008-4274

1. 一种基于 IP 网的多媒体实时授课系统 RealClass, 其特征在于: 它包括, 授课端、听课端、课堂服务中心以及多点控制单元 (Multipoint Control Unit, MCU) 四个部分组成;

其中授课端是指教师通过网络为听课端提供授课现场的实时视频、音频及电子教案, 并通过电子白板、应用程序共享以及文本 Chat 等工具与听课端进行交互, 此外授课教师还可对整个教学过程进行控制的授课平台;

授课端由一实现与课堂服务中心进行状态、控制信息的交互, 并能根据交互结果对底层的视频、音频数据采集、传输、发送、回放进行的事务管理模块, 底层的视频、音频数据处理是由 Avmeeting 控件实现; Avmeeting 控件从本地的媒体设备中读取视频、音频数据, 并通过 QoS 数据发送控件发送到本网段, 即授课端网段的 MCU; 同时调用 QoS 数据接收模块从本网段 MCU 读取焦点听课端的视频、音频数据, 并将此数据与授课端的媒体数据在本地进行回放;

听课端是指学生通过网络接收授课端的视频、音频及电子教案信息, 当听课端获得交互权限后, 还可与授课端进行视频、音频或其它方式的交互, 听课端与听课端之间在授课教师许可的情况下也可通过分组方式进行交互的听课平台;

听课端由一实现与课堂服务中心进行状态、控制信息的交互, 并能根据交互结果对底层的视频、音频数据采集、传输、发送、回放进行的事务管理模块, 底层的视频、音频数据处理同样是由 Avmeeting 控件实现; Avmeeting 控件从本地的媒体设备中读取视频、音频数据, 并通过 QoS 数据发送控件发送到授课端网段的 MCU; 同时调用 QoS 数据接收模块从本网段 MCU 读取授课端的视频、音频数据, 对于非焦点听课端, 读取的是授课端和焦点听课端融合后的数据, 并将此数据与本地的媒体数据在本地进行回放;

课堂服务中心是指管理员用于对授课系统进行监控的控制台, 并充当视频会议系统中的 GateKeeper; 其通过与授课端、听课端、MCU 的状态和控制消息交互, 实时获取 RealClass 系统的状态, 并对 RealClass 系统各个端结点的工作进行控制;

多点控制单元 (Multipoint Control Unit, MCU) 是完成系统中视频、音频数据的路由、转发以及处理, 系统采用了分布式 MCU 机制, 即在系统所分布的每个网段中都存在一个 MCU 结点;

多点控制单元 (Multipoint Control Unit, MCU) 可分为多点控制器 MC 和多点处理器 MP 两部分;

多点控制器 MC 包含多点交互控制和 QoS 参数确定两个模块; 多点交互控制模块的通过对多点处理器 MP 的操作实现对远程教学过程中的授课端和听课端的视频、音频交互控制, 此外, 它还根据与课堂服务中心交互的命令与状态信息生成当前 MCU 与其它端结点, 即授课端、听课端、其它 MCU 连接的状态信息; QoS 参数确定模块根据此状态生成当前 MCU 的各个输出媒体流的 QoS 参数, 此参数将传递给多点处理器 MP 中的 QoS 控制机制, 作为其为每个输出媒体流分配网络带宽的依据;

MP 可分为 QoS 控制机制和视频、音频数据处理两个模块;

其中视频、音频数据处理模块是 MP 的核心模块, 由 MC 中的多点交互控制模块驱动, 视频、音频数据处理模块获取 QoS 控制机制整形后的数据, 对其进行数据融合和转发等处理, 并将处理后的数据传送到 QoS 控制机制进行发送;

MP 的 QoS 控制机制根据 MC 传递来的 QoS 参数以及当前的网络状态进行 QoS 控制, 包括对视频、音频数据处理模块处理后的多媒体数据进行带宽分配和发送, 以及对由其它端结点发送来的多媒体数据进行整形。

2. 根据权利要求 1 所述的基于 IP 网的多媒体实时授课系统 RealClass, 其特征在于: 所述的 QoS 控制机制是一种基于分层结构的动态自适应 QoS 分布控制模型, 该模型一方面采用基于端结点的分布控制, 并在端结点引入动态自适应的流控机制, 使得本模型不仅具有较强的可扩展性, 而且还能根据网络状态自动调整控制策略; 另一方面, 采用多角度解决策略, 将与 QoS 控制相关的网络技术、视频编码技术以及 FEC (Forward Error Correction) 容错技术基于分层结构进行有机集成, 从而使得 QoS 模型在带宽受限的情况下也能为媒体流传输提供较好的服务质量;

对 QoS 问题的解决采用的是多角度、多层次的解决策略; 把 QoS 控制模型分为 QoS 控制层、RTP 层以及 FEC 层三个层次, 并将网络传输、视频编码以及 FEC 容错这三个 QoS 的具体控制措施有机分布到这三个层次之中;

QoS 控制层: 主要实现多媒体数据传输的准入控制、流量控制和区分服务; 主要措施为: 对音频数据流采用固定带宽分配策略; 对于视频数据流, 则采用动态自适应的带宽分配策略; 在此基础上, 根据 H.261、H.263、MPEG I、MPEG II 等视频编码的特性, 视频编码帧可以分

为设计了一个选择发送视频帧的算法,本算法保证了在数据包不丢失的情况下,发送到对方的编码数据都可被正确解码,且解码后的数据具有最优的表现质量;

RTP 层: 主要实现多媒体数据包的实时、有序传输,并通过 RTCP 包的交互动态获取当前数据传输的时延、抖动、丢包率等网络状态参数,这些参数不仅是 QoS 控制层进行带宽分配的依据,也是 FEC 层实施动态 FEC 机制的主要依据;

FEC 层: 主要功能是为上层提供一条透明的可靠传输通路,避免因网络传输丢包导致的重传;本层还引入了反馈机制,即根据当前的丢包率,动态控制 FEC 数据包的发送。

3. 根据权利要求 1 所述的基于 IP 网的多媒体实时授课系统 RealClass,其特征在于:所述的多点控制单元 MCU 是一种基于分布式 MCU 的组播数据的跨网段传输机制;MCU 分布在每个网段中;在各网段内部,MCU 向同网段的各个结点以组播方式发送数据;在各网段之间,授课端 MCU 向其它各个网段 MCU 以单播方式发送数据;此外,授课端与焦点听课端的数据则单播发送到授课端 MCU;该机制实现了广域网上视音频数据的组播传输,有效提高了网络带宽利用率,使得系统具有较好的可扩展性。

4. 根据权利要求 1 所述的基于 IP 网的多媒体实时授课系统 RealClass,其特征在于:所述的视频、音频数据采用两种课件同步录制机制:一种是实时录制工具 RealScreen,即实时采集教案内容、教学现场的视频、音频数据,并压缩生成流媒体(如 rm 或 MPEG-IV 等格式)文件,供课后点播;二是 HTML-Recorder,即将教学内容转换成超文本(HTML 文档),并和实时录制视频、音频同步集成,形成“HTML+流媒体”格式的课件;

实时录制工具 RealScreen 采用了第一种技术路线:在授课端,RealScreen 通过定时截取屏幕或窗口获得当前教案内容的内存 BMP 映像,同时将当前视频捕获设备采集到的视频帧数据转化成内存 BMP 格式,然后将两部分 BMP 数据在位图基础上进行融合,并将融合后的数据发送到流媒体压缩引擎上;RealScreen 同时还将音频设备采集到的音频数据也发送到该引擎;压缩引擎将接收到的融合数据作为视频数据进行实时编码,同时将接收到的音频数据也进行实时编码,并将编码后数据一起发送到流媒体服务器;此外,流媒体压缩引擎还可将编码数据以 RM 文件形式在本地进行保存;

HTML-Recorder 工具采用技术路线二,HTML-Recorder 工具可以和 RealClass 系统配合运行,是一个后台程序;它一方面实时地采集教师的视频、音频数据并生成 ASF 流媒体文件,另一方面它将教师对教案的操作实时记录在时间戳日志中,包括教案换页、显示每个条目的时间戳等;当运行结束时,HTML-Recorder 自动将时间戳写入 ASF 文件;并将 PowerPoint 格式的教案转化为一组 HTML 网页及相关媒体文件;最后,将携带时间戳的 ASF 文件与 HTML 格式的教案按照固定的框架(Frame)进行封装,从而生成教学课件。

基于 IP 网的多媒体实时授课系统

一、所属领域

本发明属于计算机设计与应用技术领域，涉及计算机软件、信息传递技术、多媒体技术以及网络教育/远程教育。特别涉及一种基于 IP 网的多媒体实时授课系统

二、背景技术

随着信息传递技术的发展，网络教育/远程教育已普遍被人们所接受，在这方面的研究也取得了显著的成就。申请人经过查新，检索到三篇与本发明相关的属于视频会议（Video Conference）领域的专利，它们分别是：

1. Method for video telephony over a hybrid network;
2. System, method and article of manufacture with integrated video conferencing billing in a communication system architecture;
3. System, method and article of manufacture for communications utilizing calling, plans in a hybrid network.

在上述专利 1 中，发明人提出了一套在分组交换网下开展多媒体视频会议的方法，该方法主要有以下特点：

- 1) 支持群组交互模式下多媒体数据的发送、接收；
- 2) 多媒体数据采用 RTP（Real-time Transport Protocol）协议进行传输；
- 3) 采用 RSVP（Resource reSerVation Protocol）协议对交互过程中的视音频质量进行控制；
- 4) 数据传输支持单播（unicast）、组播（multicast）、广播（broadcast）三种模式；
- 5) 交互过程可以保存为本地媒体文件；
- 6) 采用目录管理机制对会议成员进行管理。
- 7) 采用软硬件相结合的实现策略。

在专利 2 中，发明人提出了一套在视频会议系统中创建呼叫的过程，该过程主要描述了呼叫双方的认证、授权、能力协商，此外还定义了呼叫消息所携带数据的内容与格式。目前，视频会议系统一般都采用类似的呼叫机制。

在专利 3 中，发明人提出了一套在分组交换网下开展多媒体视频会议的方法，并实现了该系统。此套方法与系统主要有以下特点：

- 1) 支持多个分组间基于 RTP 协议的组播数据发送、接收；
- 2) 采用目录管理机制对会议成员进行管理；
- 3) 通过创建虚拟现实环境，实现分组间的用户相互感知，并能通过虚拟对象进行交互；

- 4) 采用目录管理机制对会议成员进行管理;
- 5) 交互过程可以保存为本地媒体文件;
- 6) 采用软硬件相结合的实现策略。

根据上述查新, 现有系统在通信方面存在以下四方面的问题:

1. 不能支持组播数据的跨网段传输, 网络带宽资源利用率较低, 进而导致系统所能支持的端节点数目以及分布范围都存在很大局限性。
2. 大多数同步实时授课系统直接采用了视频会议系统中的多点群组交互机制, 这种交互机制不适合于实时授课。采用这种交互机制, 不仅增加了上层管理的复杂程度, 同时也造成不必要的系统资源开销, 特别是网络带宽资源的开销。
3. 缺乏自适应的 QoS 控制机制, 主要表现在: 一、QoS 控制机制无法适应网络状态的动态性。二、QoS 控制机制对分布环境异构性的适应能力较差。
4. 缺乏课件自动生成机制, 只能将教学现场保存为媒体文件, 无法生成教案与教学现场同步的课件。

三、发明内容

根据上述现有技术存在的缺陷或不足, 本发明提供了一种基于 IP 网的多媒体实时授课系统 RealClass。主要包括以下几方面内容:

1. 多媒体实时授课系统 RealClass 的组成

多媒体实时授课系统 RealClass 是一个面向实时网络教学的分布式多媒体应用系统。从物理分布上可分为授课端、听课端、课堂服务中心以及多点控制单元 (Multipoint Control Unit, MCU) 四个部分。其中, 授课端是教师的控制台, 通过网络为听课端提供授课现场的实时视音频及电子教案, 并通过电子白板、应用程序共享以及文本 Chat 等工具与听课端进行交互, 此外授课教师还可对整个教学过程进行控制。听课端是学生的听课平台, 通过网络接收授课端的视频、音频及电子教案信息, 当听课端获得交互权限后, 还可与授课端进行视频、音频或其它方式的交互, 听课端与听课端之间在授课教师许可的情况下也可通过分组方式进行交互。课堂服务中心是管理员的控制台, 用于对授课系统进行监控, 并充当视频会议系统中的 GateKeeper。MCU 主要完成系统中视频、音频数据的路由、转发以及处理 (如多路视频、音频融合), 从工作机制上可分为集中式 MCU 和分布式 MCU。在本系统中, 为了扩大规模, 实现多媒体数据的自适应传输, 采用了分布式 MCU 机制, 即在系统所分布的每个网段中都存在一个 MCU 结点。

2. 组播数据跨网段传输机制

研究目的: 解决媒体数据的跨网段 (子网) 高效传输问题, 以及视频、音频融合问题。

研究背景: 在交互式多媒体同步实时授课中, 当结点数多于两个时, 需要引入称为多点

控制单元 (Multipoint Control Unit, MCU) 的实体 (Entity)。它的功能主要包括两点: ① 支持各结点的视频、音频数据的跨网段传输并进行有效控制; ② 对交互过程中的视频、音频数据进行融合处理。

目前, MCU 大多采用集中式, 由专门的 MCU 硬件实现, 成本高, 扩展性差。

本发明的解决策略:

在 RealClass 系统的设计中, 本发明提出并实现了一种基于软件的分布式 MCU 机制, 即在每个独立的网段, 设置一个 MCU, 从而形成一个级联的树型 MCU。采用单播和组播相结合的数据发送策略, 即各 MCU 之间采用单播数据发送, 各个网段内部采用组播数据发送, 这样不仅可以大大节省网络带宽, 而且还实现了组播数据的跨网段传输和视频、音频数据的融合, 有效地提高了系统的可扩展性。

3. 视音频融合机制

研究目的: 研究解决授课端和听课端视频、音频融合的问题, 将两路或多路视频、音频数据合并成一路数据, 使得融合后的视频具有“画中画”的效果, 对于音频, 则具有“混音”效果。

问题背景: 在实时教学系统中, 视频、音频融合是十分需要的。因为听课端需要感受到一个“真实”的教学环境, 因此需要把授课端及正在和授课端进行交互的“焦点”听课端的视频、音频数据进行融合, 这样不仅可以大大减少数据量和网络带宽的占用, 而且还可以为异地听课端提供一个更加形象、逼真的教学视听环境。

目前, 有关视频、音频融合的基本原理大致雷同, 本系统的特点是: 采用软件实现, 支持两路或四路视频融合, 和 MCU 有机集成, 无需附加软硬件。

4. 动态自适应 QoS 控制机制

研究目的: 在没有传输质量保障的 IP 网上, 研究解决如何保障实时教学过程的多媒体数据传输质量的问题, 即如何解决传输过程中延迟、抖动、丢包等问题。

问题背景: 现有 TCP/IP 协议采用的是尽力而为 (best-effort) 的服务机制, 该机制虽能较好地满足诸如 WWW、Email、FTP 等 Internet 的非实时应用, 但它不适合于诸如网络实时教学等的实时多媒体应用。因此, 要实现基于 IP 网的交互式同步实时授课, 必须在现有 TCP/IP 传输机制的基础上增加服务质量 (QoS, Quality of Service) 控制机制。

目前, 国际上的一些标准化组织, 如 IETF 等已提出了基于 IP 网实现 QoS 控制机制的相关协议和模型, 如 RTP/RTCP 协议、IntServ/RSVP 模型、DiffServ 模型及 MPLS 模型等。这些协议和模型为 QoS 控制机制的深层研究提供了较好的基础, 但在实际应用中都存在一定的局限性。主要表现为:

(1) 无法适应网络状态的动态性, 现有 QoS 控制机制无法根据可用带宽、丢包率等网

络状态参数的动态变化而动态调整控制策略，即不具备自适应能力。

(2) 对分布环境异构性的适应能力较差。分布环境的异构性主要表现在分布的各个结点其软硬件体系结构和数据发送方式可能存在差异。前者表现为结点可能具有不同的接入速率、主机性能，甚至不同的操作系统；后者则表现为结点可能采用组播或单播发送方式。现有 QoS 控制机制基本上不具备适应上述分布异构环境，可扩展性较差。

(3) 目前 QoS 控制机制仅局限于从网络角度提出解决策略。这种单一策略的控制机制在网络负载较重时，提供的服务质量明显下降。

本发明提出了一种基于分层结构的动态自适应 QoS 分布控制模型。其主要设计思想为：一方面，借鉴 DiffServ 模型中基于端结点的分布控制思想，并在端结点引入动态自适应的流控机制，使得本模型不仅具有可扩展性，而且还能根据网络状态自动调整控制策略。另一方面，采用多角度解决策略，将与 QoS 控制相关的网络技术、视频编码技术以及 FEC (Forward Error Correction) 容错技术等进行有机集成，从而使得本模型在带宽受限的情况下也能为媒体流传输提供合适的服务质量。

5. 多媒体教学现场同步录制技术

研究目的：同步实时录制教学现场及教案内容，生成流媒体课件，供学生课后点播学习。

问题背景：能够在实时教学过程中，一边进行实时授课，一边自动进行同步课件录制与生成，是一件十分有意义的工作。不仅能够快速简捷地生成课件资源，减少课件制作的重复工作，而且为学生课后自主点播学习提供了机会。

目前，这方面工作存在的问题是：数据量太大，生成的课件无法进行后期修改和维护，模式单一。

本发明的解决方法：

根据教学实际，本发明提出了两种解决思路：一种是实时采集教案内容、教学现场的视频、音频数据，并压缩生成流媒体（如 rm 或 MPEG-IV 等格式）文件，供课后点播；二是将教学内容转换成超文本（HTML 文档），并和实时录制视频、音频同步集成，形成“HTML+流媒体”格式的课件。这两种解决方案各有特点，可以视实际情况选择使用，主要区别在于是否对授课端计算机屏幕上的教案部分实现流式视频编码。

本发明通过教学现场的多媒体录制和网络传输，实现教学现场的直播，并通过师生多媒体多模式交互、教师自然板书授课、教学内容检索、应用程序共享、课件同步浏览、电子白板、课件实时录制、课堂管理等功能，解决传统课堂教学在时间和空间上的制约问题，大大扩展了教学规模，能够实现名师授课及教育资源的共享。

四、附图说明

图 1、本发明的多媒体实时授课系统 RealClass 的网络分布示意图；

- 图 2、本发明的实时授课系统 RealClass 的授课端工作机制示意图；
图 3、本发明的实时授课系统 RealClass 的听课端工作机制示意图；
图 4、本发明的实时授课系统 RealClass 的课堂服务中心工作机制示意图；
图 5、本发明的实时授课系统 RealClass 的 MCU 实现框架示意图；
图 6、RealClass 系统中端口使用规定；
图 7、MCU 实现的分布式 MCU 的工作机制示意图；
图 8、视频数据融合的工作机制示意图；
图 9、融合后的视频效果示意图；
图 10、本发明的音频数据融合的工作机制示意图；
图 11、基于分层结构的动态自适应 QoS 控制模型示意图；
图 12、本发明的 RealScreen 系统的工作机制示意图；
图 13、本发明的 HTML-Recorder 工具的工作机制示意图；
图 14、HTML-Recorder 工具生成的课件形式示意图；
图 15、授课端用户界面；
图 16、听课端用户界面；
图 17、课堂服务中心的用户界面。

五、具体实施方式

为了更清楚的理解本发明，以下结合附图对本发明作进一步的详细描述。

5.1 实时教学系统的组成

基于 IP 网的多媒体实时授课系统 RealClass 包括，授课端、听课端、课堂服务中心以及多点控制单元（Multipoint Control Unit, MCU）四个部分组成。

其中授课端，是指教师通过网络为听课端提供授课现场的实时视频、音频及电子教案，并通过电子白板、应用程序共享以及文本 Chat 等工具与听课端进行交互，此外授课教师还可对整个教学过程进行控制的授课平台。

授课端由一实现与课堂服务中心进行状态、控制信息的交互，并能根据交互结果对底层的视频、音频数据采集、传输、发送、回放进行的事务管理模块，底层的视频、音频数据处理是由 Avmeeting 控件实现；Avmeeting 控件从本地的媒体设备中读取视频、音频数据，并通过 QoS 数据发送控件发送到本网段，即授课端网段的 MCU；同时调用 QoS 数据接收模块从本网段 MCU 读取焦点听课端的视频、音频数据，并将此数据与授课端的媒体数据在本地进行回放。

听课端是指学生通过网络接收授课端的视频、音频及电子教案信息，当听课端获得交互权限后，还可与授课端进行视频、音频或其它方式的交互，听课端与听课端之间在授课教师

许可的情况下也可通过分组方式进行交互的听课平台。

听课端由一实现与课堂服务中心进行状态、控制信息的交互，并能根据交互结果对底层的视频、音频数据采集、传输、发送、回放进行的事务管理模块，底层的视频、音频数据处理同样是由 Avmeeting 控件实现；Avmeeting 控件从本地的媒体设备中读取视频、音频数据，并通过 QoS 数据发送控件发送到授课端网段的 MCU；同时调用 QoS 数据接收模块从本网段 MCU 读取授课端的视频、音频数据，对于非焦点听课端，读取的是授课端和焦点听课端融合后的数据，并将此数据与本地的媒体数据在本地进行回放。

课堂服务中心是指管理员用于对授课系统进行监控的控制台，并充当视频会议系统中的 GateKeeper；其通过与授课端、听课端、MCU 的状态和控制消息交互，实时获取 RealClass 系统的状态，并对 RealClass 系统各个端结点的工作进行控制。

多点控制单元（Multipoint Control Unit, MCU）是完成系统中视频、音频数据的路由、转发以及处理，系统采用了分布式 MCU 机制，即在系统所分布的每个网段中都存在一个 MCU 结点。

5.2 子系统设计

以下对 RealClass 系统中的各个子系统（包括授课端子系统、听课端子系统、课堂服务中心子系统及 MCU 子系统）的工作机制进行说明，对于各个子系统的详细设计说明参看相应的详细设计方案。

5.2.1 授课端工作机制

授课端的工作机制图 2 所示，其中授课端事务管理模块主要实现与课堂服务中心进行状态、控制信息的交互，并能根据交互结果对底层的视频、音频数据采集、传输、发送、回放进行。底层的视频、音频数据处理是由 Avmeeting 控件实现的。Avmeeting 控件从本地的媒体设备中读取视频、音频数据，并通过 QoS 数据发送控件发送到本网段（授课端网段）的 MCU；同时调用 QoS 数据接收模块从本网段 MCU 读取焦点听课端的视频、音频数据，并将此数据与授课端的媒体数据在本地进行回放。

5.2.2 听课端工作机制

听课端的工作机制图 3 所示，其中听课端事务管理模块主要实现与课堂服务中心进行状态、控制信息的交互，并能根据交互结果对底层的视频、音频数据采集、传输、发送、回放进行。底层的视频、音频数据处理同样是由 Avmeeting 控件实现的。Avmeeting 控件从本地的媒体设备中读取视频、音频数据，并通过 QoS 数据发送控件发送到授课端网段的 MCU；同时调用 QoS 数据接收模块从本网段 MCU 读取授课端的视频、音频数据（对于非焦点听课端，读取的是授课端和焦点听课端融合后的数据），并将此数据与本地的媒体数据在本地进行回放。

5.2.3 课堂服务中心工作机制

课堂服务中心工作机制如图 4 所示。其通过与授课端、听课端、MCU 的状态和控制消息交互，实时获取 RealClass 系统的状态，并对 RealClass 系统各个端结点的工作进行控制。

5.2.4 MCU 工作机制

MCU 可分为多点控制器 MC 和多点处理器 MP 两部分，其实现框架如图 5 所示。

MC 包含多点交互控制和 QoS 参数确定两个模块。多点交互控制模块的主要功能是通过 MP 的操作实现对远程教学过程中的授课端和听课端的视频、音频交互控制，此外，它还与课堂服务中心交互的命令与状态信息生成当前 MCU 与其它端结点（授课端、听课端、其它 MCU）连接的状态信息。QoS 参数确定模块根据此状态生成当前 MCU 的各个输出媒体流的 QoS 参数，此参数将传递给 MP 中的 QoS 控制机制，作为其为每个输出媒体流分配网络带宽的依据。

MP 可分为 QoS 控制机制和视频、音频数据处理两个模块。其中视频、音频数据处理模块是 MP 的核心模块，由 MC 中的多点交互控制模块驱动。视频、音频数据处理模块获取 QoS 控制机制整形后的数据，对其进行数据融合和转发等处理，并将处理后的数据传送到 QoS 控制机制进行发送。MP 的 QoS 控制机制根据 MC 传递来的 QoS 参数以及当前的网络状态进行 QoS 控制，主要包括对视频、音频数据处理模块处理后的多媒体数据进行带宽分配和发送，以及对由其它端结点发送来的多媒体数据进行整形。

5.2.5 各个模块间的接口

由 MCU 实现框架图，MCU 的内部接口包括以下几个部分：

■ MC 中的 QoS 参数确定模块与多点交互控制模块之间的接口

MC 中的 QoS 参数确定模块与多点交互控制模块之间的接口是多点交互控制模块根据由课堂服务中心发送来的全局 MCU 状态、本网段端结点状态生成的全局 MCU 状态列表和本网段端结点状态列表。

■ MC 中的 QoS 参数确定模块与 MP 中的 QoS 控制机制之间的接口

MC 中的 QoS 参数确定模块与 MP 中的 QoS 控制机制之间的接口是 QoS 参数确定模块根据全局 MCU 状态列表和本网段端结点状态列表生成的各个输出媒体流的 QoS 参数，即媒体流的带宽许可范围。

■ MC 中的多点交互控制模块与 MP 中的视音频数据处理模块之间的接口

MC 中的多点交互控制模块与 MP 中的视音频数据处理模块之间的接口是由 MC 生成的或转发课堂服务中心的视音频数据融合命令、视音频数据转发命令。

■ MP 中的 QoS 控制机制与视音频数据处理模块之间的接口

MP 中的 QoS 控制机制与视音频数据处理模块之间的接口包括经 QoS 控制机制整形

后传输到视音频数据处理模块的数据，以及经视音频数据处理模块的处理后传输到 QoS 控制机制的数据。

由 MCU 与 RealClass 系统中的其他端结点的交互关系，MCU 的外部接口包括以下几个部分：

■ MC 中的多点交互控制模块与课堂服务中心之间的接口

MC 中的多点交互控制模块与课堂服务中心之间的接口主要是两者之间交互的命令和状态，包括：

- (1) 由 MCU 向课堂服务中心发送的命令和状态，主要有 MCU 登录命令、MCU 退出命令、端结点视音频控制命令。
- (2) 由课堂服务中心向 MCU 发送的命令和状态，主要有全局 MCU 状态、本网段端结点状态、视音频数据融合命令、视音频数据转发命令。

■ MP 中的 QoS 控制机制与授课端和听课端的接口

MP 中的 QoS 控制机制与授课端、听课端、其它 MCU 的接口主要指他们之间视音频数据收发端口的规定，对端口合理的规定，可以大大简化交互的控制逻辑。

在 RealClass 系统中，规定授课端向 2001 端口发送视频数据，向 2000 端口发送音频数据，从 4001 端口接收视频数据，从 4000 端口接收音频数据；听课端向 3001 端口发送视频数据，向 3000 端口发送音频数据，从 4001 端口接收视频数据，从 4000 端口接收音频数据；MCU 从 2000~2001 端口获取读取授课端视音频数据，从 3000~3001 端口获取读取听课端视音频数据，处理后向 4000~4001 端口发送视音频数据。如图 6 示。

■ MP 中的 QoS 控制机制与其他 MCU 中的 QoS 控制机制之间的接口

MP 中的 QoS 控制机制与其他 MCU 中的 QoS 控制机制之间的接口即 MP 之间的接口。在 RealClass 系统中，规定 MP 之间进行视音频数据交互时，视频数据使用 5001 端口，音频数据使用 5000 端口。如图 6 示。

5.3 主要关键技术

5.3.1 组播数据跨网段传输机制

由于 IP 网一般不支持组播数据的跨网段发送，在本系统中，本发明基于分布式 MCU 实现了组播数据的跨网段，其原理如图 7 所示。主要思想是采用级联的 MCU 对组播数据进行转发。具体工作机制如下：

- I. 授课端将本地视频、音频数据单播发送到授课端 MCU，并接收经授课端 MCU 组播发送的数据，此数据是授课端视频、音频与焦点听课端视频、音频融合后的数据。
- II. 授课端 MCU 将接收到的授课端视频、音频数据与焦点听课端视频、音频数据融合后，单播发送到各个非授课端 MCU 上，并组播到本地的授课端与各个听课端。

III. 非授课端 MCU 将来自授课端 MCU 的视频、音频数据组播到本地各个听课端。

IV. 焦点听课端将本地视频、音频数据单播发送到授课端 MCU。

上述机制使得同步实时授课系统的负载不再受限于听课端的结点数目，而仅受限于听课端所分布的网段数目，从而有效地提高了系统的可扩展性。

在交互过程中，本发明提出了一种适合实时教学的多媒体群组交互模式——镜头焦点交互模式，其含义为：在同一时刻，允许一个称为“焦点”的听课端结点与教师进行交互，并将两者的多媒体数据经 MCU 融合后发送到其余各个结点，从而实现实时同步教学过程的传输。在这一过程中，学生可以随时向教师提出交互请求，教师可根据需要随时切换焦点。采用这种镜头焦点交互模式，不仅简化了实时授课系统的上层管理，而且还避免了不必要的带宽资源开销。

5.3.2 视频、音频数据融合技术

本发明的解决方法：

■ 视频融合

当 MCU 中的 MP 接收到两路视频数据后，对其进行解码，解码后的数据为 CIF 格式或 BMP 格式，若解码后为 BMP 格式，则首先将其转化为 CIF 格式。对于两路 CIF 格式的数据，首先将其中一路视频帧转化成一路 QCIF 格式，然后再和另一路视频帧进行融合，根据视频的原编码算法，再对融合后的视频帧进行编码。视频融合的工作机制如图 8 所示。融合后的视频效果如图 9 所示。

■ 音频融合

当 MP 接收到两路音频数据后，对其进行解码，解码后的数据为线性的音频数据，然后将两路线性的音频数据进行代数迭加，进而根据音频的原编码算法，再对迭加后的音频数据进行编码。音频融合的工作机制如图 10 所示。

5.3.3 动态自适应 QoS 控制机制

本 QoS 控制模型的实现要点为：

■ 基于端结点的分布控制机制

本模型引入了具有自适应能力的流控机制，它能根据当前网络状态为视频数据流动态分配网络带宽，同时对每个视频数据包能否进入网络进行准入判断。网络状态主要包括可用带宽、传输延迟以及丢包率。该机制所依据的网络状态是端结点间的整体状态，这充分掩盖了内部结点的异构性，使得该模型能较好地适用于异构的分布环境。

■ 基于分层结构的多角度解决策略

和以往不同的是：本发明对 QoS 问题的解决采用的是多角度、多层次的解决策略。把

QoS 控制模型分为 QoS 控制层、RTP 层以及 FEC 层三个层次,如图 11 所示,并将网络传输、视频编码以及 FEC 容错这三个 QoS 的具体控制措施有机分布到这三个层次之中。

QoS 控制层: 主要实现多媒体数据传输的准入控制、流量控制和区分服务。主要措施为:对音频数据流采用固定带宽分配策略;对于视频数据流,则采用动态自适应的带宽分配策略,在此基础上,根据每个视频数据包的编码特征确定其是否可以进入网络,即是否发送到 RTP 层。

RTP 层: 主要实现多媒体数据包的实时、有序传输,并通过 RTCP 包的交互动态获取当前数据传输的时延、抖动、丢包率等网络状态参数,这些参数不仅是 QoS 控制层进行带宽分配的依据,也是 FEC 层实施动态 FEC 机制的主要依据。

FEC 层: 主要功能是为上层提供一条透明的可靠传输通路,避免因网络传输丢包导致的重传。本层还引入了反馈机制,即根据当前的丢包率,动态控制 FEC 数据包的发送。

上述三个层次之间有明确的接口,每一层通过其接口向上层提供透明的服务。在发送端,媒体数据由应用程序经 QoS 控制层、RTP 层以及 FEC 层逐层分解并添加包头信息后发送到网络系统;在接收端,媒体数据由网络系统经 FEC 层、RTP 层以及 QoS 控制层逐层还原后返回到应用程序。

■ QoS 控制层

QoS 控制层的主要功能是对输出媒体流进行流量限制。根据多媒体应用对视音频数据 QoS 需求特点,本模型中,采用音频数据优先的策略,基于带宽预留思想为音频数据流分配固定带宽;对于视频数据,则采用动态自适应的带宽分配机制,并在此基础上,引入基于视频编码特征的信包调度模块,使得当网络负载较重时,只发送对 QoS 影响较大的媒体数据。以下对视频数据的带宽分配机制与信包调度机制进行说明,并给出调度算法。

视频数据的带宽分配采用了“线性增长成倍衰减”的策略,此策略为:若当前的数据包丢失率小于某个阈值,则线性增加分配带宽值;若当前的数据包丢失率大于某个阈值,则将当前带宽值乘上衰减比例因子,作为新的分配带宽值。基于上述策略的带宽分配机制既能充分利用带宽资源,又能有效地降低丢包率。

视频数据的信包调度由令牌管理模块和数据包发送判定模块协同完成。令牌管理模块的工作机制是基于“漏桶算法”^[2],其主要作用是对媒体数据发送速率进行限制,使得发送速率符合带宽分配模块所确定的速率。该子模块以带宽分配模块所确定的速率生成令牌,如令牌池中令牌已满,则令牌停止生成;当数据包发送判定子模块确定要发送一个数据包时,首先通知该子模块减去数据包大小的令牌数。数据包发送判定子模块根据当前令牌池中的令牌数以及当前视频数据包的编码特征判定是否将数据包发送到网络。根据 H.261、H.263、MPEG I、MPEG II 等视频编码的特性,视频编码帧可以分为 I 帧(帧内编码)、P 帧(帧间编码)以

及 B 帧（双向预测）^[4,5]。考虑到 I 帧、P 帧、B 帧解码过程中的相互依赖关系，传输视频数据必须遵循以下原则：

- (1) 在传输过程中，I 帧优先级最高，P 帧优先级次之，B 帧优先级最低；
- (2) 传输 P 帧的前提是必须传输与之关联的 I 帧或 P 帧；
- (3) 传输 B 帧的前提是必须传输与之前后相邻的 I 帧或 P 帧。

基于上述三个规则，提出如下调度算法对视频数据包的发送进行判定，此算法的主要思想是在带宽受限的情况下，尽可能多地发送优先级高的数据包。调度算法描述如下：

```

case 数据包类型 of
I 帧: if B 帧队列非空 and
当前令牌数 > (当前数据包大小 + B 帧队列大小) then
发送 B 帧数据包和当前数据包并将令牌数减去此次发送数据包大小之和
else
if 当前令牌数 > 当前数据包大小 then
发送当前数据包并将令牌数减去当前数据包大小；
清空 B 帧队列；
P 帧: if B 帧队列非空 and
当前令牌数 > (当前数据包大小 + B 帧队列大小 + 平均关键帧大小) then
发送 B 帧数据包和当前数据包并将令牌数减去此次发送 数据包大小之和；
else
if 此帧关联的 I 帧已发送 and
当前令牌数 > (当前数据包大小 + 平均关键帧大小) then
发送当前数据包并将令牌数减去当前数据包大小；
清空 B 帧队列；
B 帧: if 此帧的前一个 I 帧或 P 帧已发送 then
将此帧放入队列；
end.

```

其中，平均关键帧大小可通过对多个关键帧的大小统计获得，B 帧队列大小指当前队列中所有 B 帧大小之和。

若在信包调度过程中不考虑数据包的编码特性，则在接收端接收到数据后，很可能因为其相关的 I 帧数据或 P 帧数据没有发送而无法解码，成为无用数据，这使得信道的实际利用率大大降低。本算法保证了在数据包不丢失的情况下，发送到对方的编码数据都可被正确解码，且解码后的数据具有最优的表现质量。

■ 3.2 FEC 层

FEC 层的主要功能是基于 FEC 容错技术为上层提供一条的“较可靠”的传输通道, FEC 层虽无法提供完全可靠的数据传输, 但能有效地降低丢包率。其工作机理是: 在发送端把 k 个原始数据包编码生成 n ($n > k$) 个数据包, 使得 n 中的任何 k 个数据包都能够恢复出原始数据包, 这样在接收端只要接收到任意 m ($m > k-1$) 个数据包, 就可以恢复原始的 k 个数据包, 即允许传输过程最多丢失 $n-k$ 个数据包^[3, 6]。相对于目前网络传输中常用的 ARQ (Automatic Repeat Request) 控制机制, FEC 机制主要有以下优点:

- (1) FEC 机制是由数据发送端实现的, 不需要与数据接收端进行交互, 实现机制较 ARQ 简单;
- (2) 适用于组播或广播模式的数据发送, 在这种一对多的发送模式下, 性能不会随着接收方的数目增加而下降, 即具有较好的可扩展性^[3]。

FEC 机制的主要缺点就是在网络丢包率比较低的情况下, 仍然生成固定数目的冗余数据包, 这些数据包不仅占用了一定网络带宽, 而且也增大了传输延迟^[3]。

本模型的 FEC 层将 FEC 机制和反馈机制结合起来, 实现了动态自适应的 FEC 机制。该机制不仅能有效地降低数据传输过程中的丢包率, 而且在丢包率较低的情况下能有效减少冗余数据占用的带宽。其主要机理为:

- (1) 根据媒体流丢包率的许可范围, 确定 n 、 k 值;
- (2) 采用公式①确定实际发送的数据包数 m ($k \leq m \leq n$), 此数值应保证在当前丢包率情况下, 接收方仍能以较高的概率恢复编码前的 k 个数据包;
- (3) 编码器顺序将 k 个数据包编码成 n 个数据包;
- (4) 随机选择 m 个数据包发送。

在动态 FEC 机制中, m 的确定是其中的关键。设当前采用 $\langle n, k \rangle$ 编码, 当前丢包率为 r , 允许丢包率为 R ($R < r$), m 应满足下述公式:

$$\sum_{i=0}^{m-k} C_m^i (1-r)^{m-i} r^i > (1-R)^k \quad \text{①}$$

其中, 上式的左边表示当前丢包率为 r 时, 接收端接收到 m 中的 k 个以上 (包括 k 个) 数据包的概率, 即能还原出 k 个数据包的概率; 上式的右边表示在允许丢包率为 R 时, 接收端正确接收到 k 个连续数据包的概率, 显然, 只有满足上述不等式关系时, 才有必要进行 FEC 编码。采用枚举法, 逐个选取 m 值 ($k \leq m \leq n$) 代入上式测试, 最后选取能使上式成立的最小 m 值。当 $r < R$ 或 $r \gg R$ 时, 可能不存在使上述公式成立的 m 值, 则当前没有必要进行 FEC 编码, FEC 层直接将 RTP 数据包发送到网络系统。因而, 上述公式不仅用于计算 FEC 编码

传输中应实际发送的数据包数,而且也可用于判断是否对 RTP 数据包进行 FEC 编码。

5.3.4 多媒体教学现场同步录制技术

■ 流媒体课件实时录制工具——RealScreen

RealScreen 工具采用的是技术路线一,它不仅可以实现实时教学的流媒体课件录制,而且还可以实现教学现场的直播,RealScreen 由授课端、服务器端以及听课端三个部分组成。其中授课端主要实现了教学现场与教案内容的录制与编码;服务器端主要实现流媒体课件的发布与点播服务,听课端实现课件的回放。RealScreen 的工作原理如图 12 所示。

在授课端,RealScreen 通过定时截取屏幕或窗口获得当前教案内容的内存 BMP 映像,同时将当前视频捕获设备采集到的视频帧数据转化成内存 BMP 格式,然后将两部分 BMP 数据在位图基础上进行融合,并将融合后的数据发送到流媒体压缩引擎上;RealScreen 同时还将音频设备采集到的音频数据也发送到该引擎。压缩引擎将接收到的融合数据作为视频数据进行实时编码,同时将接收到的音频数据也进行实时编码,并将编码后数据一起发送到流媒体服务器。此外,流媒体压缩引擎还可将编码数据以 RM 文件形式在本地进行保存。在服务器端,流媒体服务器将接收到的数据以 RM 文件形式进行保存,同时根据听课端的请求,采用 RTSP/RTP 协议将流媒体数据传输到听课端。在听课端,用户采用 RealPlayer 将接收到的编码数据进行解码、回放。

RealScreen 将授课教案和视频、音频进行有机融合,这使得 RealScreen 不仅可以支持任何形式的教案,自动、同步、实时地生成流媒体课件,而且还可以实现教学现场直播和按需点播的同步进行。(如图 12 所示)这种方式的缺点是数据量较大,对课件的后期改进比较困难。

■ 基于“HTML+流媒体”的实时课件录制工具——HTML-Recorder

HTML-Recorder 工具采用技术路线二,能够较好地解决了 RealScreen 工具存在的数据量大、带宽占用多、后期修改困难等问题,其工作原理如图 13 示。

HTML-Recorder 工具可以和 RealClass 系统配合运行,是一个后台程序。它一方面实时地采集教师的视频、音频数据并生成 ASF 流媒体文件,另一方面它将教师对教案的操作实时记录在时间戳日志中,包括教案换页、显示每个条目的时间戳等。当运行结束时,HTML-Recorder 自动将时间戳写入 ASF 文件;并将 PowerPoint 格式的教案转化为一系列 HTML 网页及相关媒体文件;最后,将携带时间戳的 ASF 文件与 HTML 格式的教师教案按照固定的框架(Frame)进行封装,从而生成教学课件。界面形式如图 14 示。

与 RealScreen 工具相比,HTML-Recorder 工具生成的课件在传输时只占用相对较小的带宽,并具有更好视频、音频质量,而且生成的课件可以进行后期修改,但缺点是目前课件的格式只支持 Powerpoint 格式。

RealScreen 工具和 HTML-Recorder 工具的有机结合, 可以提供完善的课件录制功能。

5.3.5 界面设计

界面是系统功能和特点的集中反映, 为了真正实现交互式同步实时授课的效果, 需要将授课端现场、课件内容及其它必要的辅助功能以多窗口的形式展现在教师或学生的面前。

■ 授课端用户界面

授课端用户界面如图 15 示。包括授课端视频显示区、听课端视频显示区、课件内容浏览区、功能按钮区、听课端状态显示区等 5 部分。其中听课端视频显示区可以由教师任意切换到任一听课端。

■ 听课端用户界面

听课端用户界面如图 16 示。包括听课端视频显示区、授课端视频显示区、课件浏览区、功能按钮区、授课端状态显示区等 5 部分, 其中, 课件浏览区和授课端显示区这两部分与授课端保持一致。

■ 课堂服务中心的用户界面

课堂服务中心的用户界面如图 17 示。包括听课端视频显示区、授课端视频显示区、授课端状态显示区、听课端状态显示区、功能按钮区以及 MCU 状态信息显示区等 6 部分。其中, 听课端视频显示区可以任意切换到任何一个听课点上。

本发明与现有技术相比, 所产生的效果是:

1. 听课端个数: 可同时支持 20 个以上教室的同步实时授课, 并可同时支持 4 个以上不同子网之间的视音频交互。
2. 课件格式: 支持 HTML/XML 文档、Word 文档、Powerpoint、MPEG4、RM、RAM、VRML 等格式的课件, 并能方便地扩充新的格式。
3. 课件质量: 以 640×480 窗口或 800×600 全屏方式清晰显示课件内容。
4. 视频交互: 遵循 H.261 标准实现视频数据实时采集和回放、融合, 视频质量由 QoS 控制机制保证。
5. 语音交互: 遵循 G.711 标准实现音频数据实时采集和回放、融合。
6. 视音频同步: 唇音同步误差≤0.5 秒。
7. 系统延时: 系统的影像、声音及课件内容在 Internet/校园网上传输时延小于 2 秒。
8. 应用程序共享: 遵循 T.120 协议, 支持各种 Windows 应用程序共享。
9. 电子白板: 遵循 T.120 协议。

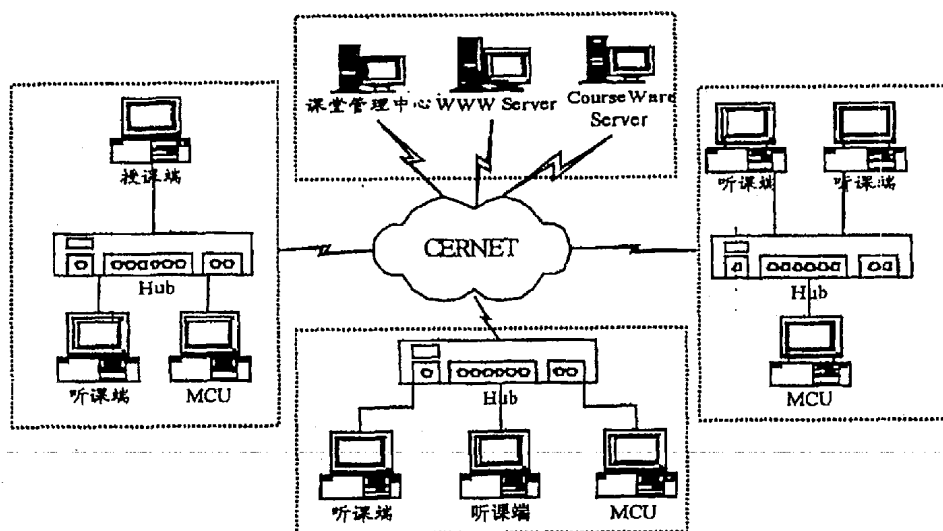


图 1

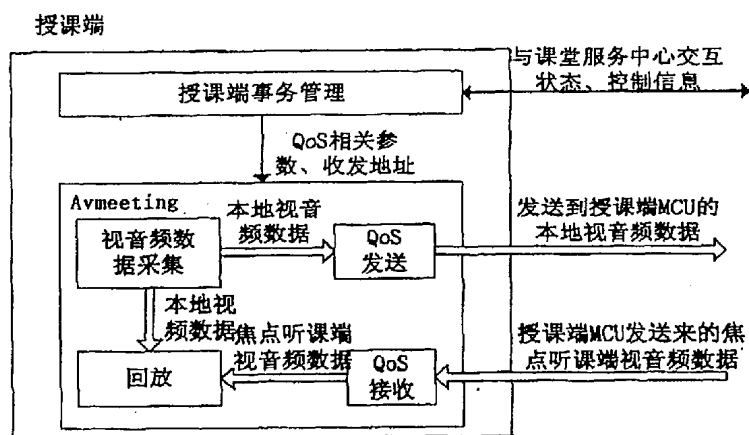


图 2

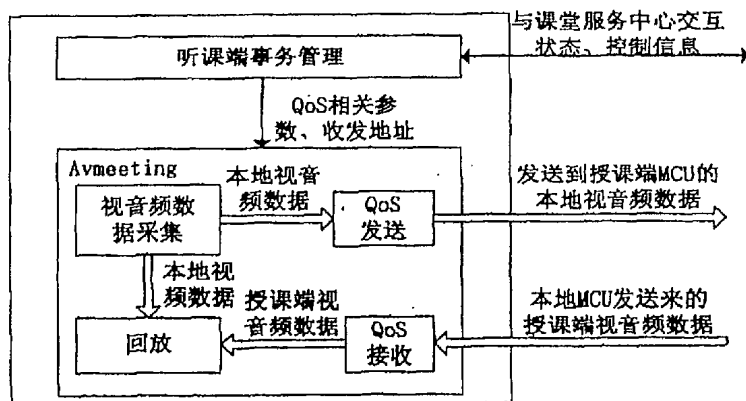


图 3

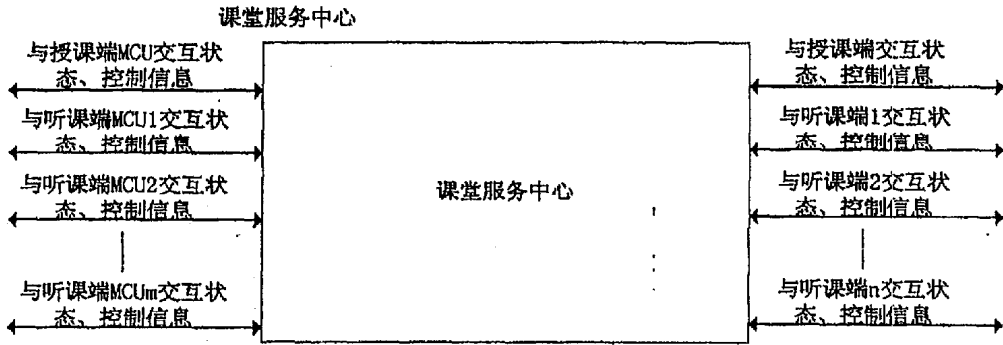


图 4

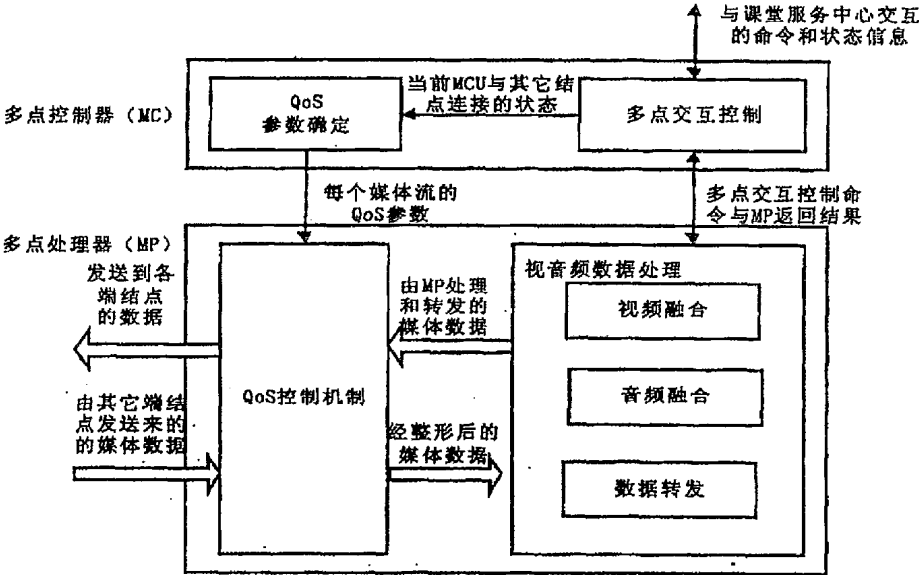


图 5

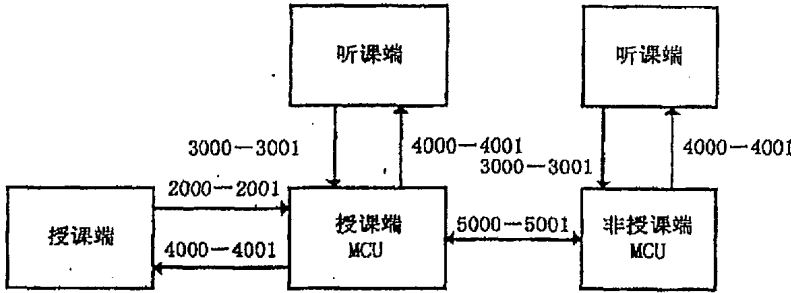


图 6

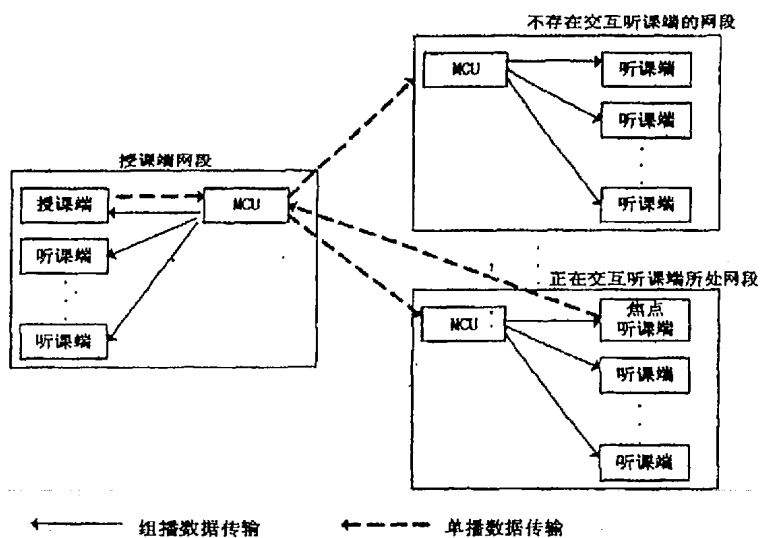


图 7

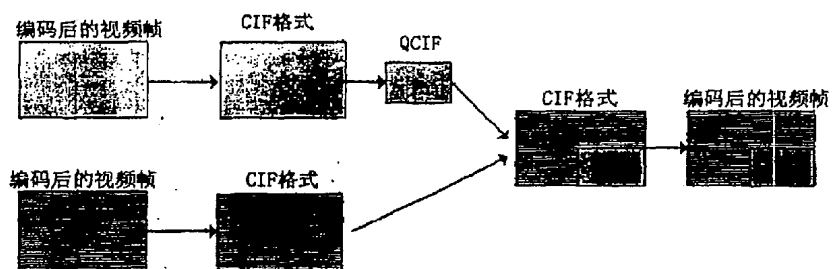


图 8



图 9

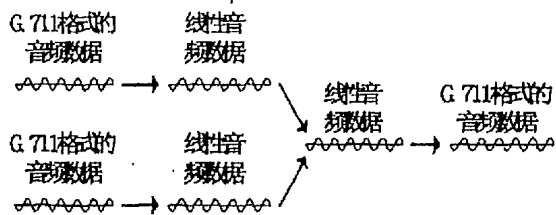


图 10

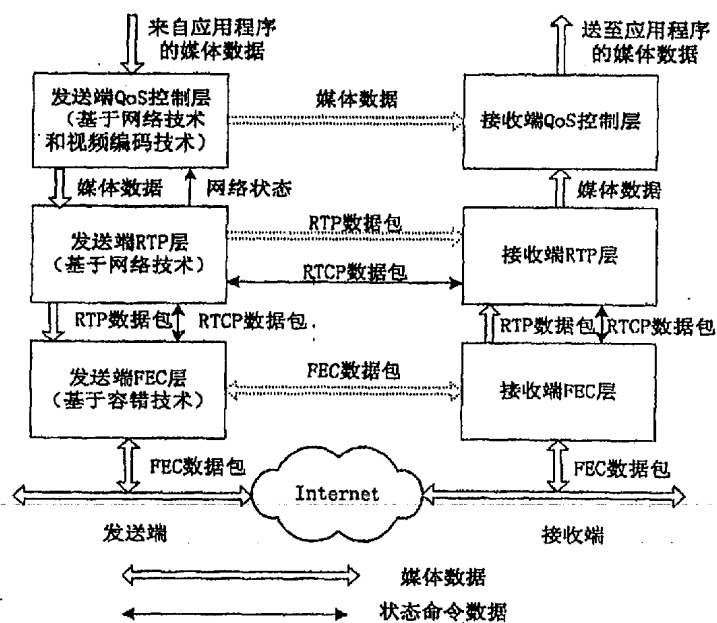


图 11

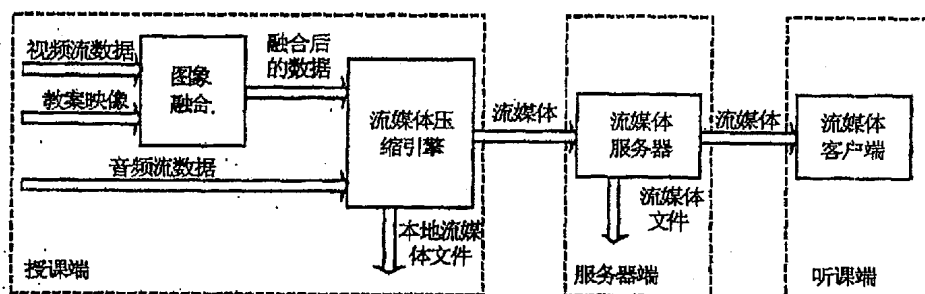


图 12

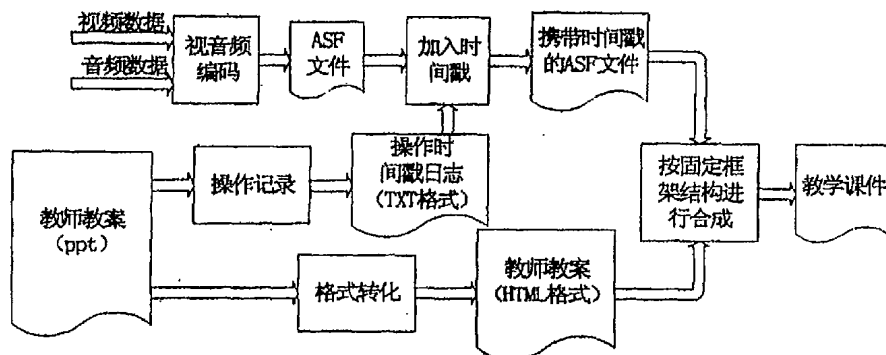


图 13

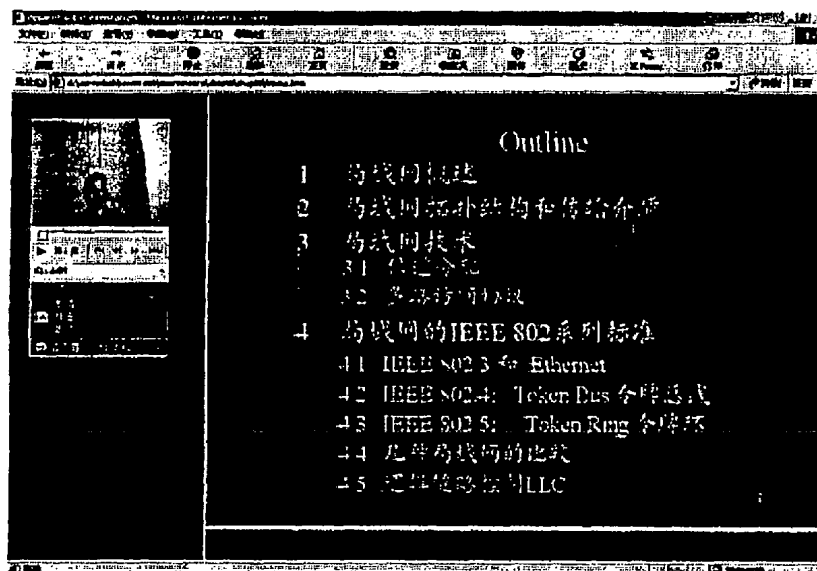


图 14

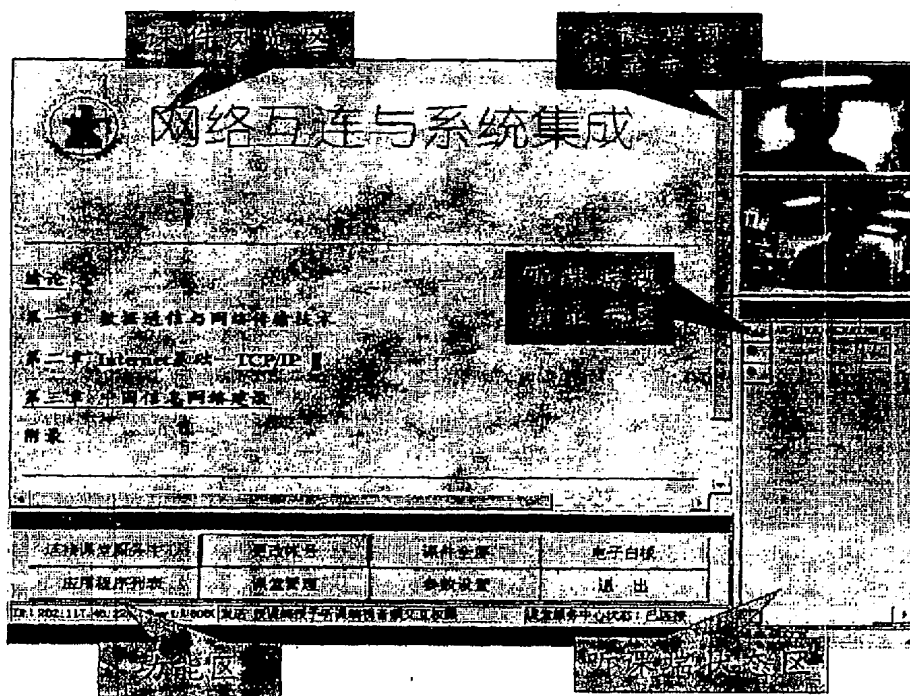


图 15

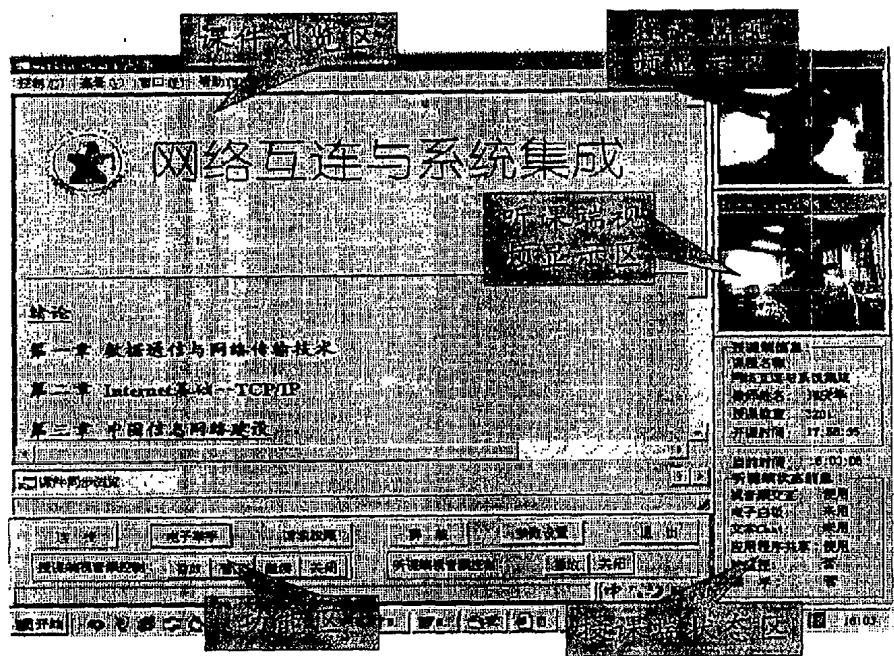


图 16

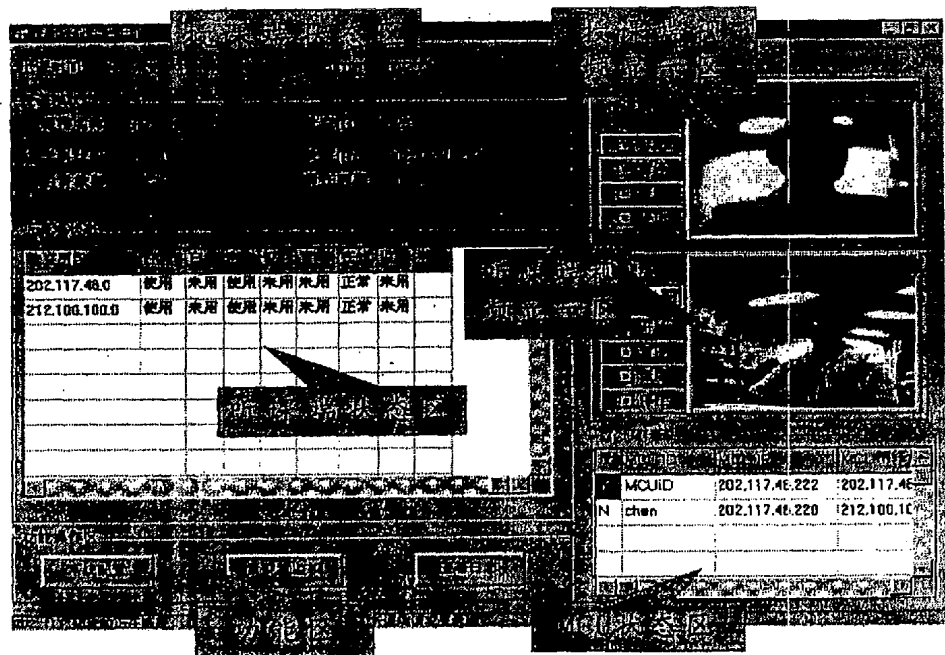


图 17